

Design and manufacture of synthetic pinnæ for studying head-related transfer functions (HRTFs)

Eric Michael Sumner, Morris Riedel & Runar Unnthorsson

To cite this article: Eric Michael Sumner, Morris Riedel & Runar Unnthorsson (2025) Design and manufacture of synthetic pinnæ for studying head-related transfer functions (HRTFs), Cogent Engineering, 12:1, 2536150, DOI: [10.1080/23311916.2025.2536150](https://doi.org/10.1080/23311916.2025.2536150)

To link to this article: <https://doi.org/10.1080/23311916.2025.2536150>



© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 15 Aug 2025.



Submit your article to this journal [↗](#)



Article views: 282



View related articles [↗](#)



View Crossmark data [↗](#)

Design and manufacture of synthetic pinnae for studying head-related transfer functions (HRTFs)

Eric Michael Sumner^a , Morris Riedel^a  and Runar Unnthorsson^b 

^aComputer Science, University of Iceland, Reykjavík, Iceland; ^bMechanical Engineering, University of Iceland, Reykjavík, Iceland

ABSTRACT

A major obstacle in investigating the relationship between the anthropometric properties of an individual's pinnae and their corresponding head-related transfer function (HRTF) is the inability to perform controlled experiments free from other confounding factors. We propose a method for performing these experiments by manufacturing a digitally-altered replica pinna. The process involves 3D-scanning a real pinna, performing the desired alterations, manufacturing the replica, and measuring the acoustic properties of the replica. Here, we describe our design and manufacturing methodology, 3D printing a mold of the altered pinna, which is then used to make a silicone casting, as well as our acoustic measurement apparatus and procedures. To demonstrate the effectiveness of our process, we manufactured an unaltered pinna replica and show that the replica has acoustic properties comparable to the original within the range of human hearing (≤ 20 kHz).

ARTICLE HISTORY

Received 4 September 2024
Revised 4 February 2025
Accepted 18 February 2025

KEYWORDS

Head-related transfer functions (HRTFs); spatial audio; additive manufacturing; anthropometric variables; experimental design

SUBJECTS

CAD CAE CAM – Computing and Information Technology; Mechanical Engineering Design; Acoustical Engineering

1. Introduction

One of the most important aspects of human hearing is the sense of sound localization, the ability to determine where a sound originates. The ability to stimulate this sense artificially has the potential to improve users' immersion and situational awareness. Although improved acoustic fidelity has a comparable effect on the user's immersion as improvements in visual and tracking fidelity, most virtual reality work to date has been in these latter areas, with relatively little attention paid to presenting an accurate soundscape (Cummings & Bailenson, 2016); incorporating individualized acoustic rendering will further enhance the feeling of presence in the virtual environment. In an augmented reality context, it can enhance safety by directing the user's attention to something outside their current field of vision. This enhancement has been successfully applied in some military applications (Bronkhorst, 1995), but wider adoption is blocked in part by a lack of an economical solution for obtaining individualized head-related transfer functions (HRTFs).

In addition to the enhancement of virtual and augmented reality systems used for job training (Ajmera & Gonen, 2020), individualized auditory displays have a wide variety of applications within the healthcare sector. In bandwidth-limited telehealth environments, spatial auditory displays are a more effective use of limited bandwidth than visual displays (Liu & Meng, 2007). An accurate characterization of how an individual's pinnae affect the sound waves that reach their ear canal will also allow in-ear assistive devices for the hearing-impaired to identify the direction of incoming sounds and alert the wearer to dangers outside their current field of vision (Denk et al., 2018). Adding auditory feedback

CONTACT Eric Michael Sumner  ems36@hi.is  Computer Science, University of Iceland, Reykjavík, Iceland

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

systems improves the engagement of stroke patients in neurorehabilitation activities (Rosati et al., 2011). Spagnol et al. incorporated personalized spatial audio displays into electronic travel aids (ETAs) (Spagnol, Wersényi, et al., 2018), resulting in an effective wayfinding system for visually-impaired individuals (Spagnol, Hoffmann, et al., 2018). Visually-impaired individuals can also navigate by way of echolocation, and spatial audio technology has enabled the development of a safe training environment that contains no physical obstructions (Andreasen et al., 2019).

The primary cues used to locate sound sources in the horizontal plane are the interaural time delay (ITD) for low frequencies (Rayleigh, 1909) and the interaural level difference (ILD) for high frequencies (Ferree & Collins, 1911). For sound sources outside the horizontal plane, the auditory system relies on interference patterns generated by internal reflections within the pinna (Hebrank & Wright, 1974). In particular, for any given source direction, some frequencies have a node located at the entrance to the ear canal, producing a comb filter. The effects of this comb filter on incoming sounds are interpreted to determine the direction of the sound source relative to the listener (Baumgartner et al., 2014).

Given the far-field acoustic transfer function of a person's body, known as their HRTF, it is possible to accurately simulate sounds coming from any particular direction or combination of directions (Morimoto & Ando, 1980). This reduces the problem to one of obtaining the relevant transfer function. Owing to natural anatomical variation, each person has a unique HRTF; attempting to use a generalized HRTF produces inaccurate results (Bronkhorst, 1995). Individualized HRTFs are therefore desirable for all applications and are necessary for some (Jenny & Reuter, 2020).

Directly measuring a person's individualized HRTF is costly. It requires access to an anechoic chamber, specialized measurement equipment, and a significant amount of time. Developing a method to accurately estimate an individualized HRTF using less costly means is currently a field of active research. Most estimation approaches fall along one of two lines, either finding a good match from a database of HRTFs via listening tests (Pelzer et al., 2020) or synthesizing an HRTF based on anthropometric measurements (Grijalva et al., 2016). Both approaches have shown promising results, but they are still significantly weaker than either directly-measured HRTFs or real-world localization performance.

At the Acoustic and Tactile Engineering (ACUTE) laboratory, we are pursuing this second approach for generating synthetic HRTFs. As a part of this work, we would like to study the effects of particular anthropometric properties in isolation. Unfortunately, the shape of human pinnæ has too many varying properties to be able to construct a controlled series. While it may be possible to find a set of people that form a good sample of one property, such as the distance between tragus and anti-tragus (see Figure 1 for definition of terms), the effect of this property will likely be obscured by the effects of the many other anthropometric variations that will necessarily be present.

Many researchers have likewise published studies attempting to isolate the effect of particular pinna microfeatures on HRTFs with some success. Those that studied data from human subjects (Algazi et al., 2001; Spagnol et al., 2013) consistently report that the confounding factors of other features limit the power of their findings. Other researchers have taken a different approach of using acoustic simulations to calculate the HRTFs of wholly synthetic pinnæ (Katz, 2001; Ghorbal et al., 2017; Guezenoc & Séguier, 2020). These studies do not suffer from the confounding effects, but suffer slightly in showing that their simulated results actually reflect physical reality. The approach presented here of manufacturing and taking practical HRTF measurements of digitally-specified pinnæ is designed to bridge the gap between these two approaches. It will enable the cross-checking of simulated results via physical experiment, and its own accuracy can be verified by reproducing HRTFs of exemplar pinnæ. Though 3D printing of human pinnæ has been extensively explored in the field of reconstructive surgery (Humphries et al., 2022), we are unaware of any other efforts to apply these techniques to HRTF research.

In this study, we aim to validate this approach by replicating a commercial pinna insert and demonstrating that no acoustical anomalies are introduced during the process. Section 2 describes our method of scanning, manufacturing, and measuring the HRTF of an artificial pinna. Section 3 then describes the experiments we performed to validate the reliability and accuracy of our method, and Section 4 discusses the implications of these results on future studies that rely on this method and equipment.

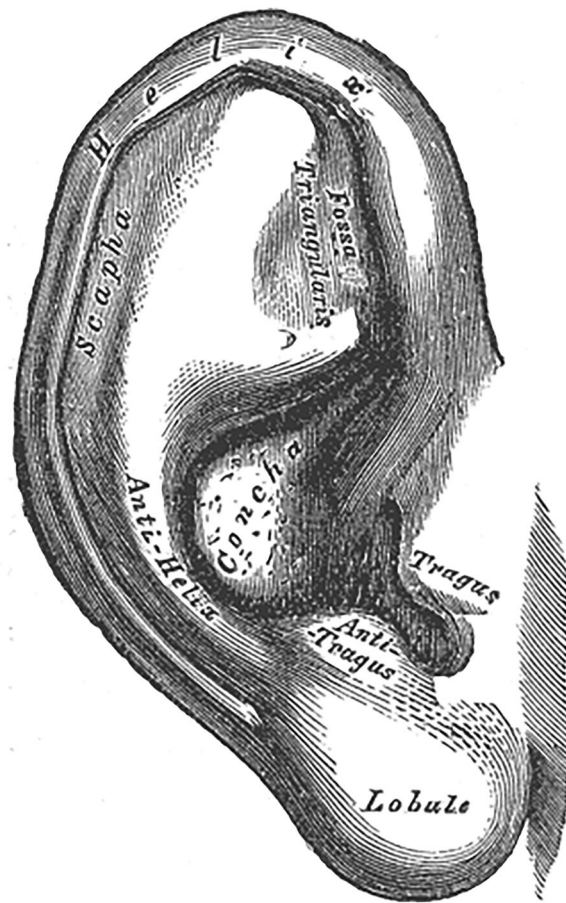


Figure 1. Anatomy of the human pinna. Reproduced from H. Gray and Lewis (1918), Plate 904.

2. Methodology

Our process, summarized in [Figure 2](#), consists of generating a 3D model of a real pinna, manufacturing a silicone insert of that model for a Knowles Electronic Manikin for Acoustic Research (KEMAR) (Burkhard & Sachs, 1975) and measuring the HRTF of the KEMAR with the insert installed. In future studies, we plan to also manufacture and measure the HRTF of synthetic pinna models, produced through digital manipulation of these scanned pinnæ. [Section 2.1](#) describes our process of scanning a pinna and producing the 3D model, [Section 2.2](#) then describes how we manufacture a silicone insert in the shape of that model, and [Section 2.3](#) describes how we measure HRTFs.

2.1. 3D scanning and modelling

We start the process by scanning the pinnæ of volunteers with a Creaform Academia 50 handheld optical scanner, depicted in [Figure 3a](#), which provides a measurement accuracy of 0.25 mm. Owing to fundamental limitations of the scanner technology, it is impossible to scan every surface of the pinna. The scanner requires a standoff distance of 400 mm from the target, which means that there are 23 degrees of parallax between the emitter and primary detector. As both of these must be able to see the target surface at the same time to take a reading, narrow features present a particular problem: When a concave feature's depth is more than 240% of its width, the edges of the feature will always occlude either the emitter or the detector from the deepest surfaces.

When scanning the pinnæ of live people, there are a couple of additional limiting factors to this scanning process. Certain surface textures, in particular hair, do not reflect a clear enough image of the excitation pattern to produce a reading. Also, the overall duration of the scan is limited based on the subject's ability to hold still. Unintentional motions will inevitably deform the subject's face enough for

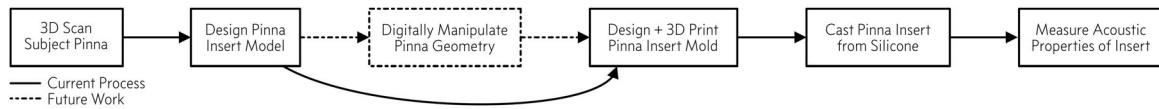


Figure 2. Process overview.

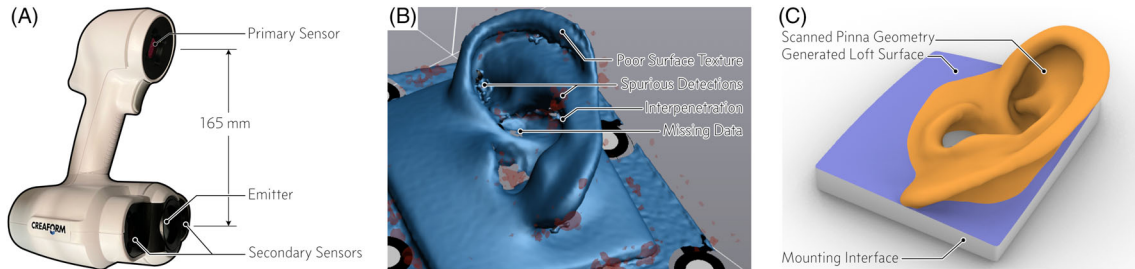


Figure 3. Modelling process. (A) Creaform Academia 50 3D scanner, (B) Raw scan example with annotated defects, (C) Final model.

the scanner software to lose its tracking such that it cannot be recovered. We mitigate this somewhat by using the scanner software's semi-rigid tracking mode which automatically corrects for small deformations that occur during the scan.

For the scan used in this study, we manufactured a scanning base to mount the commercial pinna insert on. This base can be seen in the background of Figure 3b and features retro-reflective targets to enhance the scanner's tracking algorithm. This allowed the scanner operator to pause the scan and freely reposition the object without losing tracking. Thus, the operator was able to see intermediate scan results while the scan was in process to determine which areas were still unscanned, resulting in the maximum possible scan coverage. In addition to the tracking mode, the scanning software only offers two configuration parameters: The scan resolution, which was maximized at 1 mm, and the shutter speed, which was left in automatic mode.

The resulting holes in the scanned mesh must be filled with synthetic data, as shown in Figure 3b. In addition to these holes, the raw scans also exhibit other defects, such as spurious detections, rough surface texture, or interpenetration by the back face of thin solids. These defects are corrected manually with the aid of proprietary software that is provided by the scanner's manufacturer. First, triangles representing free-floating surfaces and unnatural protrusions are deleted. Holes are then filled in high-curvature mode, which ensures that the position, tangent, and curvature of the generated surface match the scanned data at the boundary. Finally, the software's smoothing tool is applied locally to any surfaces which appear unnaturally rough.

The cleaned scan is then sent to a traditional CAD program, Rhino3D, which is used to create the final pinna insert model, shown in Figure 3c. As part of this modeling process, the scanned data is converted to a subdivision surface using Rhino3D's QuadRemesh command with a target edge length of 1 mm, which produces an additional smoothing effect. This model is composed of three distinct parts: The cleaned-up scan data (orange), mounting geometry (white), and a lofted surface (blue) to blend the two together. The mounting geometry is designed to attach to KEMAR via an interference fit around the microphone housing. The loft surface is tangent-matched to the scan data and position-matched to the interface geometry, which results in a smooth transition between the side of the KEMAR head and the scanned geometry.

2.2. Manufacturing

As the positive geometry is cast out of silicone, it is necessary to design a mold that will produce the positive geometry described above. There are several practical concerns incorporated in the design of this mold, the first of which is the ability to remove the casting from the mold once the silicone is cured. Because silicone is quite flexible, most of the undercut details of the pinna do not cause a

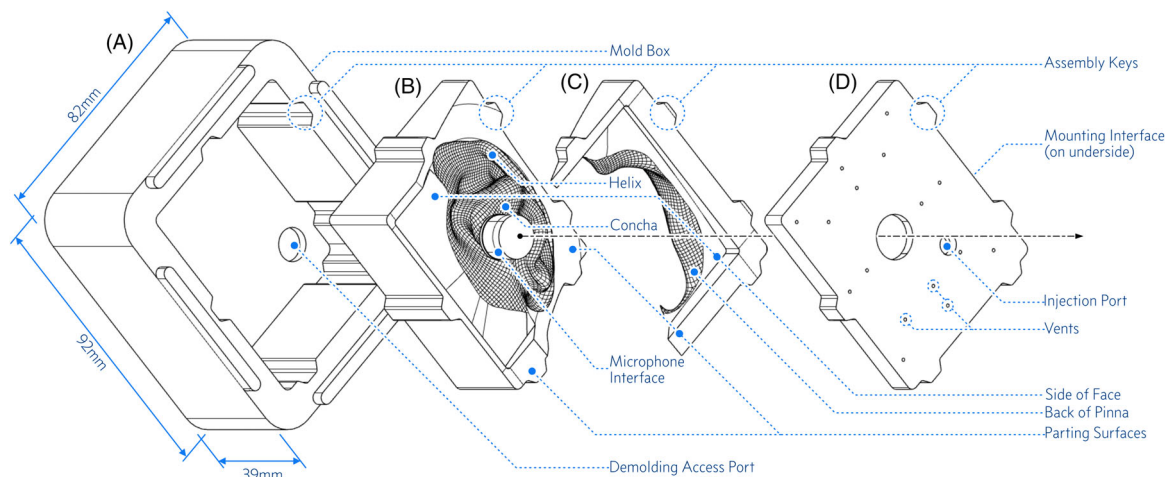


Figure 4. Multi-part mold design: (A) Mold box, (B) Pinna inner surface, (C) Pinna back surface, (D) KEMAR mounting geometry.

problem. The combination of a relatively thick helix surrounding a thin scapha, however, requires special consideration: Attempting to pull the cast helix through the scapha void risks tearing the silicone, thus destroying the casting.

2.2.1. Mold design

To address this problem, the mold has been designed in three parts, as shown in Figure 4: one that defines the detail of the interior of the pinna (b), a second that fills the void between the pinna and the face (c), and a third that defines the interface geometry on the back side of the insert (d). These are generated by constructing parting surfaces around the helix, the interface perimeter, and the microphone mount, and then subtracting the resulting mesh from the mold blank. This blank has been designed to fit inside a mold box (a) with asymmetric rails to ensure proper alignment during the mold assembly.

The other major factor to consider when designing the mold is providing a sufficient pathway for the displaced air to escape, so that the final casting is free of internal voids. Three primary escape routes exist in the design presented here:

- Along the helix parting surface and then up the perimeter of the mold box,
- Through the parting surface around the microphone mount, and
- Through vents added to the backside of the interface.

These routes were selected such that the resulting flashing can be easily removed without disturbing the most critical parts of the pinna geometry. The pinna is also oriented vertically in the mold so that any air pockets that do form will naturally migrate away from the primary surface of the pinna towards the mounting geometry, which is a less acoustically-critical region.

Each of the four mold parts has been designed with a planar face perpendicular to the mold assembly direction; for printing, this face is oriented on the platen as shown in Figure 5a. This minimizes the amount of support material necessary for the printing process. It also helps to prevent fitment problems with the final assembly. Because all of the mold parts are printed within a common coordinate system, any miscalibration of one of the printer's axes will affect all parts uniformly.

The vents are cylindrical holes with a nominal diameter of 1 mm. We verify that the vents in the printed part are clear using a set of pin gauges; the actual diameter of each vent in the example analyzed below is between 0.7 mm and 1.0 mm. In order to minimize print failures due to their small feature size, they have been oriented perpendicular to the platen. If they were placed in any other orientation, there are two likely failure modes: If support material is not used, the bridging filament on the high side of the vent could droop during printing, causing the vent to be blocked. The alternative of filling the vent with soluble support material is also problematic as the small cross-sectional area would make it difficult for the solvent to fully enter and dissolve the support, which will also result in a clogged vent.

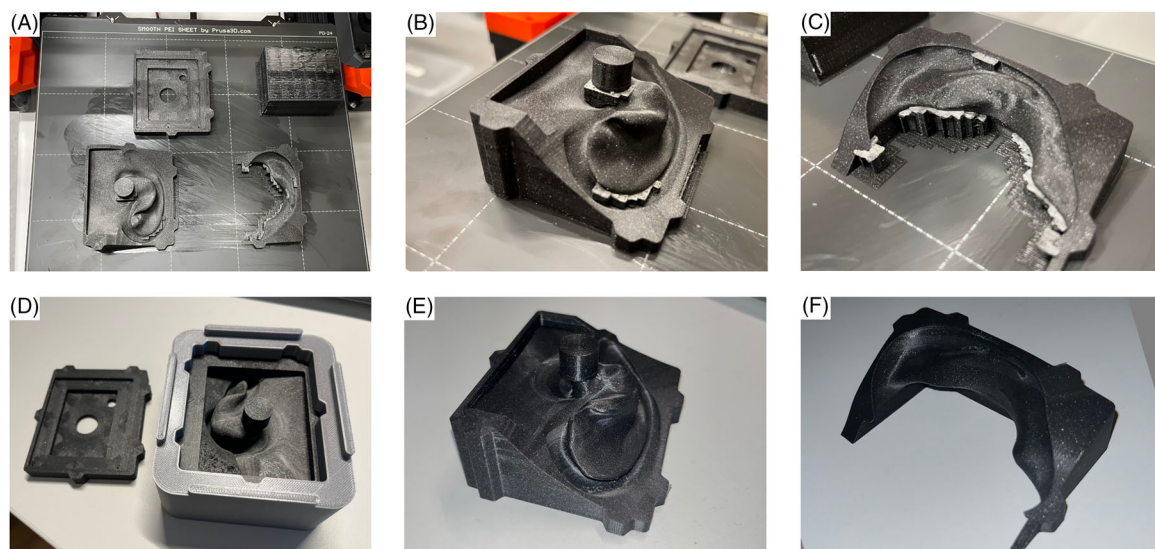


Figure 5. 3D printing results: (A) Print configuration overview, (B) Pinna inner surface detail, with supports, (C) Pinna back surface detail, with supports, (D) Partially-assembled mold, (E) Pinna inner surface detail, supports removed, (F) Pinna back surface detail, supports removed.

Table 1. Summary of print parameters^a.

	Primary Filament	Support Filament
Material	PLA	PVA
Layer thickness	0.05 mm	0.05 mm
Wall thickness	2.65 mm ^b	n/a
Vertical shell thick.	1mm	n/a
Infill	20% Gyroid	Hexagonal
Flow rate (max)	17.59 mm ³ /s	6.28 mm ³ /s
Print speed	45–140 mm/s	50–120 mm/s
Extrusion width	0.45 mm	0.36 mm
Nozzle diameter	0.4 mm	0.4 mm
Print head temp.	220 °C	200 °C
Platen temp.	60 °C	60 °C

^aParameters not listed are unchanged from manufacturer's defaults. ^bWhere possible. Thinner features are printed as solid material.

2.2.2. 3D printing

The mold and mold box are printed in polylactic acid (PLA) using a Prusa MK4S multi-material fused-deposition modelling (FDM) printer with the parameters listed in Table 1. We selected a gyroid infill pattern in order to minimize the internal stresses caused by overlapping filament within the infill pattern. Polyvinyl alcohol (PVA) is used as a support material to ensure the geometric accuracy of the various overhangs in the mold parts. After printing, the mold parts are soaked in a water bath for several hours to soften the PVA, which is then removed with a stiff brush. Figure 5 shows the result of the 3D printing process both before and after support removal. Aside from removing the support material, no post-processing needs to be performed on the printed parts.

In the course of developing this procedure, we encountered a number of practical issues during the 3D printing process, mostly relating to the support material. The earliest attempts were without support material at all, relying on the printer's ability to produce unsupported bridges; this almost worked, but resulted in an unacceptable accuracy of overhanging surfaces. It also occasionally produced a drooping loop of material that would become embedded in the casting if not manually corrected. Printing with solid support material was also problematic due to the confined spaces in which some support material needed to be placed. There was no satisfactory way to reach the location with the necessary removal tools.

The PVA soluble support material itself has also been the source of several printing issues, as it has proved to be quite susceptible to degradation from ambient humidity if improperly stored. This

degradation has manifested in several different failure modes, the two most common being jamming of the printer's material-change mechanism and the failure of deposited support material mid-print, causing the part to move and future layers to be misaligned. These problems were mitigated by storing the PVA filament in a sealed container, accompanied by dessicant; only small quantities, significantly less than an entire spool, of the PVA filament are removed from storage and installed on the printer at any given time.

Other 3D printing issues manifested as part failures during the later processes of support removal and silicone casting, which can induce significant physical stresses. In some instances, print layers delaminated during the water bath used to remove support material, presumably due to the PVA expanding as it absorbed the water. The root cause of this problem proved to be a miscalibration of the required flow rate for the PLA filament, leaving internal defects from an insufficient quantity being laid down on each layer. During the casting process, some liquid silicone inevitably finds its way between the main mold pieces and the mold box, as this is one of the gas escape routes. Overcoming the resulting adhesion during the demolding process requires a significant amount of force; if the wall thickness and/or infill ratio is reduced, the mold itself can fail instead.

2.2.3. Silicone casting

Once the parts have been thoroughly dried, the mold is assembled and held together with tape. The tape is also used to seal the holes in the bottom of the mold box, which are included to aid the demolding process. A 2-part silicone with a nominal hardness of Shore OO50 (ASTM D2240-00, 2017) is injected into the mold manually using a 60 cm³ syringe via the injection port at the top (on the back side of the KEMAR interface). We found that the silicone has sufficiently weak adhesion to the PLA print material that there was no need to add any release agent to the mold surface. After solidification, some flashing remains along the mold parting surfaces and vents, which is manually removed.

The most common failure mode for the casting and solidification process is entrapped gasses leaving voids in the final pinna, and we made several adjustments to our casting process during development to mitigate this issue:

- We tried several different commercial silicone formulations, and discovered that the key properties for a successful casting in this application are a low viscosity, which allows the liquid silicone to more easily flow into the fine details of the mold, and an extended curing time, which allows the entrapped gasses more opportunity to travel upwards, away from the acoustically-critical regions.
- It is also critical to keep the injection flow rate low, so that the boundary layer inside the mold remains approximately planar during the injection process. A fast flow rate can cause a local rise underneath the injection port, which may prematurely cut off gas escape routes and result in a void in the final product.
- The sharp edges of the mold's print layers can serve as nucleation sites that hold small gas bubbles during the injection process; to mitigate this problem, we interrupt the injection process several times to vigorously agitate the mold (i.e. bang it on the table) in order to break these gas pockets loose from the surface.
- Once the mold has been fully injected with silicone, it is held under 4 bars of pressure for solidification which notably reduces the volume of any remaining internal voids.

2.3. Acoustic measurements

The measurements are performed in an anechoic chamber, which has been certified for free-field acoustic measurements (ISO 26101, 2012) for frequencies above 200 Hz. The test stand installed in this chamber, depicted in Figure 6a, consists of a KEMAR mounted on a rotating column and a boom arm with an attached loudspeaker.

The rotational position of the column, θ , represents the azimuth angle of the emitter to the right of the KEMAR median plane. The boom arm rotates about a horizontal axis passing through the centroid of the KEMAR head. The rotational position of the boom arm, ϕ , represents the emitter's elevation angle above the horizontal plane. Due to mechanical constraints, the range of motion of the elevation axis is

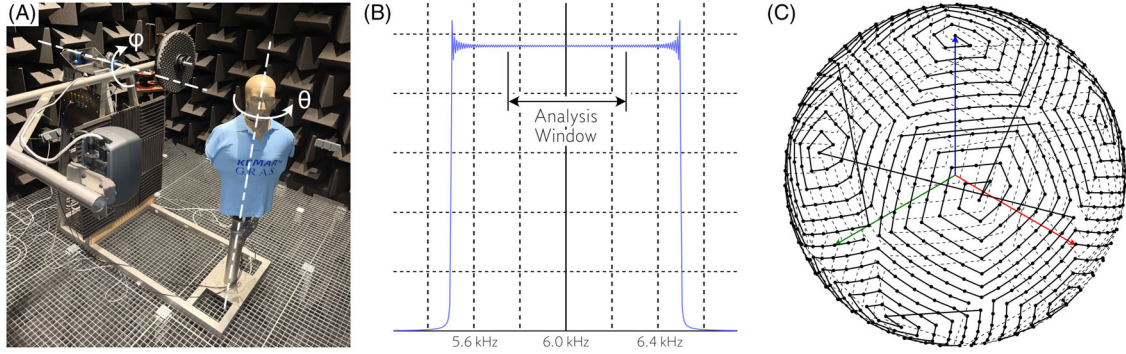


Figure 6. Acoustic measurements. (A) Photo of test stand, annotated with axes of rotation for azimuth, θ , and elevation, ϕ , coordinates. (B) Frequency-domain representation of a time-windowed 6 kHz $\pm$ 500 Hz excitation wavelet. (C) Isometric view (above right) of the spatial scan pattern; the lines indicate the measurement sequence, starting with $\phi = 90^\circ$; the axes shown represent $\theta = 0^\circ, \phi = 0^\circ$ (red), $\theta = 90^\circ, \phi = 0^\circ$ (green), and $\phi = 90^\circ$ (blue).

$-65^\circ \leq \phi \leq 90^\circ$. The elevation and azimuth axes are under computer control, driven by stepper motors with 1 degree steps. The absolute position of these axes is recorded by rotary encoders with an accuracy of 0.2° , providing a closed-loop control system.

A Genelec 8020C speaker is attached to the boom such that, at $\theta = 0, \phi = 0$, both elements lie on the median plane, and the high-frequency element is one meter away from the KEMAR and at the same height as the boom's axis of rotation; the effective elevation of the low-frequency element is 5.4 degrees higher. The cutoff frequency of the low-frequency element is 3 kHz; as the most relevant spectral information is above 4 kHz, this element should have a minimal effect on the measurements. Similar 2-driver Genelec speakers have previously been used for HRTF measurements at several research centers, and this configuration is within the bounds of the standard operating practices that they have established (Bolaños & Pulkki, 2012; F, 2019).

In prior machine learning experiments, the non-uniform density of a rectangular HRTF sampling pattern, especially around the poles, has been shown to significantly bias the resulting models (Sumner et al., 2022). Therefore, we elected to use a novel sampling grid with a relatively uniform distribution over the unit sphere. The $\{3, 5+\}_{(k,k)}$ geodesic polyhedron family (Wenninger, 1979) was chosen for this purpose because of its straightforward construction and ability to adjust the overall density by altering the parameter k . We are currently taking measurements with $k = 6$, which results in 1036 different configurations after removing the low-elevation configurations which are outside the range of motion of the test stand. The resulting sampling grid is shown in Figure 6c in isometric view.

In order to minimize the effects of any ambient noise, we use a time-division approach for taking the acoustic measurements: The final reported transfer function is a piece-wise assembly of several band-limited measurements. The excitation signal for these individual measurements is the sinc wavelet described in Eq. 1 with center frequency f_0 and bandwidth Δf . This formulation of the sinc function can interpolate between the unit impulse function at $\Delta f = 0$ and a pure sinusoid at $\Delta f = \infty$; in all cases, the frequency-domain representation is a rectangular function centered at f_0 . This interpolation represents a tradeoff between the time and frequency domains. At one end of the scale, the signal power is concentrated at a single frequency but is spread evenly across all time; at the other end, the power is concentrated at a single point in time but spread across the entire frequency spectrum. By choosing an appropriate intermediate value, a physically-realizable time-domain signal that efficiently probes a desired portion of the frequency domain can be selected.

$$\begin{aligned} \mathcal{F} \left[\cos(2\pi f_0 t) \cdot \frac{\sin(\pi \Delta f t)}{\pi \Delta f t} \right] \\ \propto \begin{cases} 1, & \omega \in [\pm 2\pi f_0 - \pi \Delta f, \pm 2\pi f_0 + \pi \Delta f] \\ 0, & \text{otherwise} \end{cases} \end{aligned} \quad (1)$$

We use a wavelet bandwidth of $\Delta f = 1\text{kHz}$, which contains 99.9% of its power in our 200 ms time slice. This is sampled at 192 kHz to minimize the possibility of high-frequency noise being aliased into the measurement band. We use only data from the central 500 Hz of each wavelet in order to minimize

the effect of the distortion from Gibbs's phenomenon (Bocher, 1906) which occurs near discontinuities in the frequency domain, as shown in Figure 6b. Gibbs's phenomenon produces a slight distortion within this region as well, which we compensate for by dividing the received frequency spectrum by the frequency spectrum of a corresponding time widow of the excitation signal; this is equivalent to performing a deconvolution operation in the time domain. The overall excitation signal is a train of 38 wavelets with center frequencies $f \in [1\text{kHz}, 20\text{kHz}]$; to minimize interference, the center frequencies of adjacent wavelets differ by at least 3kHz.

3. Results

Two experiments were conducted to validate the efficacy of the proposed method. The first, described in Section 3.1, investigates the accuracy and repeatability of our measurement equipment and the second, described in Section 3.2, investigates the fidelity of our scanning and manufacturing processes by replicating a commercial KEMAR pinna insert.

3.1. Measurement apparatus validation

We used a pair of commercially produced pinna inserts, GRAS models KB1090 and KB1091, as our control. The KEMAR acoustic response with these inserts was measured five times. The signal-to-noise ratio (SNR) of the left channel from each individual (frequency, position) pair was then estimated using the mean squared over variance method. Figure 7 shows two different projections of this SNR estimate. In Figure 7a, all the directional measurements for each frequency are grouped together; the line shows the median SNR, and the shaded regions show the range of the middle 50% and 80% SNR estimates. Across all frequencies, the 10th percentile SNR was above 7.7 dB and the median SNR was above 16 dB.

There was a notable decrease in the SNR between 4 and 8 kHz. For these frequencies, the signal wavelength is approximately the same size as the pinna. This represents a transition region where the KEMAR head starts to effectively block acoustic propagation, and pinna geometry effects begin to play a primary role. As we wish to study these pinna effects specifically, we calculated the high-band ($f \geq 8\text{kHz}$) SNR for each measurement; these are shown in Figure 7b. The horizontal axis represents the azimuth of the measurement, and the color represents the elevation. Within this band, all measurements had an SNR of at least 18 dB, with the exception of low-elevation reading where the emitter was on the far side of the KEMAR. The measurements obtained in this adverse situation remained above 12 dB.

3.2. Manufacturing process validation

In order to validate our manufacturing process, we produced a replica of the GRAS left pinna insert; Figure 8 shows our analysis of the accuracy of this replica, with a visual comparison in Figure 8a. The replica has a rougher surface texture, caused by the layer thickness of to the 3d printing process, and it is made of a slightly more pliable material. This extra pliability explains the apparent extra width in the

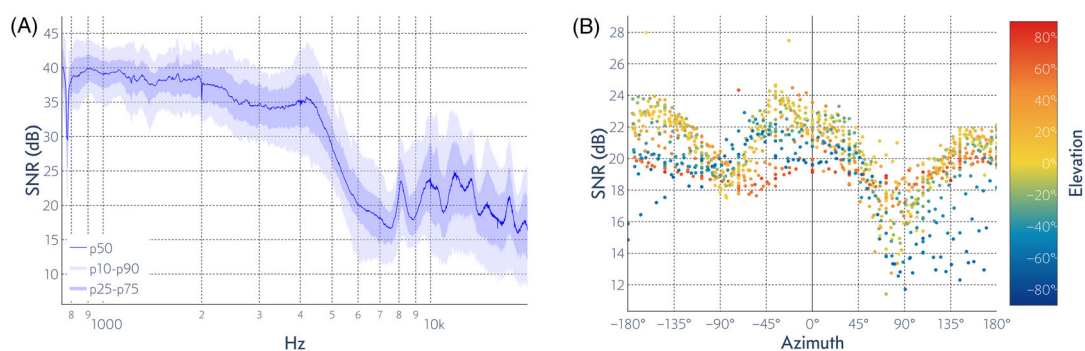


Figure 7. Measurement repeatability testing. (a) Calculated signal-to-noise ratio by frequency band, (b) High-frequency ($f \geq 8\text{kHz}$) signal-to-noise ratio by measurement direction.

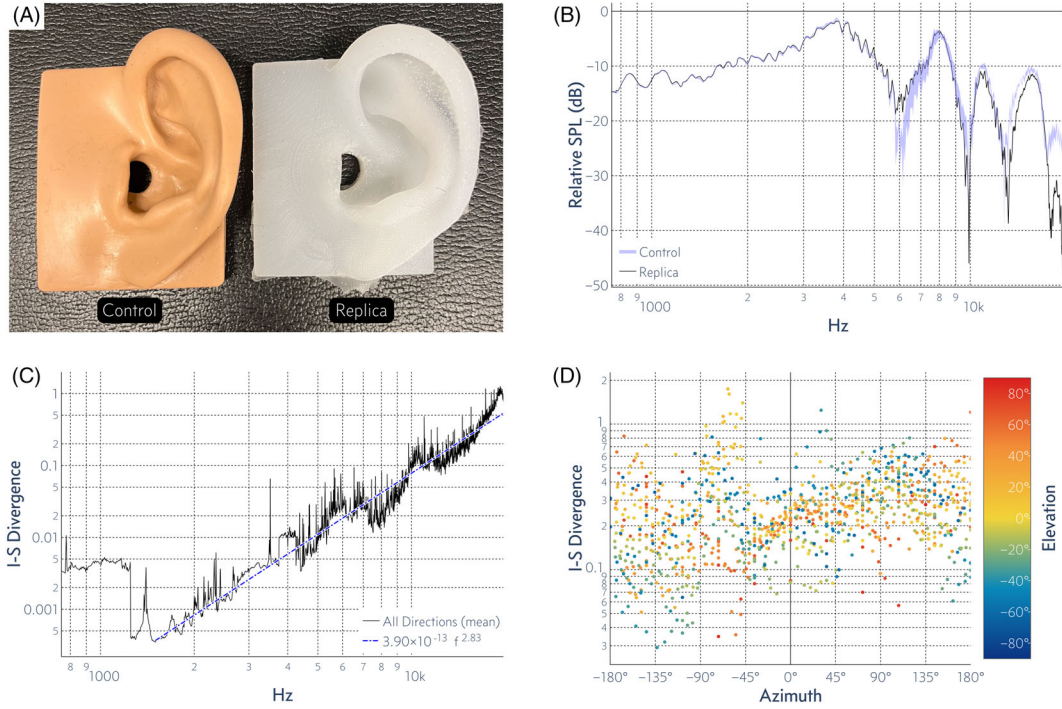


Figure 8. Comparison of control and replica pinna inserts. (a) Visual comparison, (b) Representative transfer function ($\theta = -86^\circ$, $\phi = -58^\circ$), (c) Spectral distortion of replica by frequency, (d) High-frequency ($f \geq 8\text{kHz}$) spectral distortion of replica by measurement direction.

visual comparison: With the inserts sitting flat on a table, gravitational deformation causes a slightly different angle to be presented to the camera.

Using the data from the experiment described in Section 3.1 as a control, we replaced the left pinna with our replica; Figure 8b shows a representative measurement from the replica's HRTF, recorded from a sound source that is mostly in-line with the left pinna and from a low elevation. The line shows the measurement taken from the replica, and the blue shaded area shows a $\pm 1\sigma$ confidence interval derived from the control measurements. Below 5500 Hz, the replica and the control match each other quite well. Between 5500 Hz and 17 kHz, the notch frequencies and overall shape are a good match, but there is some notable deviation in overall amplitude. Over 17 kHz, the replica's response is 10 dB below the control's.

Other directions show a similar pattern, with low frequencies tracking the control very well and high frequencies deviating somewhat while retaining a similar overall shape and qualitative features. To characterize these deviations, we estimated the Itakura-Saito divergence (R. Gray et al., 1980) d between the left channel of this measurement and the mean left channel from the apparatus validation measurements, according to Eq. 2.

$$d = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{X(\omega)}{\hat{X}(\omega)} + \log \frac{X(\omega)}{\hat{X}(\omega)} - 1 \right) d\omega \quad (2)$$

$$\approx \frac{1}{N} \sum_{i=1}^N \left(\frac{X_i}{\hat{X}_i} + \log \frac{X_i}{\hat{X}_i} - 1 \right)$$

This describes a perceptual distortion measure between a true frequency-domain signal $X(\omega)$ and its reconstruction $\hat{X}(\omega)$. For our approximation, we take N uniformly-distributed samples over a region of interest; X_i refers to the i^{th} sample in the discretized representation of X . The measurement shown in Figure 8b has a divergence of 0.61, which is higher than that of 95% of all other measurements recorded in this experiment.

Figure 8c shows the calculated divergence for each frequency sampled over the entire spatial distribution. Above 1500 Hz, there is an exponential relationship between the frequency and the distortion with a power of 2.83. At our maximum analyzed frequency of 20 kHz, this represents a divergence of

1.0. [Figure 8d](#) shows the high-band ($f \geq 8\text{kHz}$) divergence for each spatial measurement. There is little correlation between the elevation angle and the amount of signal distortion, but the measurement quality is slightly better for negative azimuths, where the left side of the KEMAR (which has the pinna under test) is facing the emitter.

4. Discussion

This study did not characterize or attempt to correct for the spectral response of either the speaker or the microphones used. However, these effects will be consistent for any data collected with the same equipment. Therefore, future studies need to take care to either only compare measurements taken on the same equipment, as has been done here, or normalize the raw data for the equipment's response using a procedure similar to that described by Spagnol et al. when collecting an HRTF database with earlier versions of this equipment (Spagnol et al., 2020, 2021).

The increasing errors at higher frequencies are mildly concerning, and further study is required to determine the underlying cause of this distortion. Some potential candidates are the stepped surface texture produced by the FDM printing process, errors introduced by 3D scanning and manual cleanup, and the different hardnesses of the pinna insert materials.

We plan to use data collected using the method presented here to train machine learning models which will be able to predict HRTFs. Anticipated challenges in this process were the main driver behind our selection of a geodesic sampling grid, but this choice also provided some advantages in the analysis presented here. In particular, the uniform sampling density of the geodesic grid allowed us to simplify our analysis. Had we used a more-traditional rectangular grid, there would have been a dense cluster of data samples around the 90° elevation direction, directly above the KEMAR. This would, in turn, have biased the aggregate measures in our analysis towards effects that occurred in that direction. In order to present a fair analysis under such circumstances, we would have needed to devise and apply some kind of corrective factor which would have complicated the analysis considerably.

In the immediate future, we plan to conduct a series of studies to investigate how this process can be made more efficient without degrading the acoustic results. It is difficult, for example, to obtain reliable 3D scans of the geometry on the rear side of the pinna. Replicating it accurately also poses significant manufacturing challenges, such as requiring deep undercuts in the mold design. If, as we suspect, these undercuts can be largely filled in without significantly affecting the acoustic results, the reliability of every step of our process can be improved: Larger defects in the 3D scan of this difficult area can be tolerated, the mold design can be simplified from three parts to two, and any entrapped gasses in the casting can more reliably migrate away from the acoustically-sensitive geometry.

Similarly, the 3D printing parameters themselves can be optimized further. The parameters presented here represent the highest-fidelity reproduction that our equipment is able to produce. It is likely that the process efficiency can be increased by adjusting the parameters such that the printing time is reduced in exchange for less accurate geometry. Increasing the print layer thickness, for example, would reduce both the printing time and the proportion of print failures due to miscalibration or other environmental issues, but it would also degrade the high-frequency acoustic results. Investigating this trade-off is a promising direction for future studies.

Overall, it is clear that this manufacturing and measurement process is a useful tool for studying HRTFs: The introduced errors are not biased for or against any particular sound source direction; even at the highest distortion levels, all of the qualitative features of an HRTF remain, in particular, the overall envelope shape and the frequencies of the comb filter notches agree between the control and experimental pinnae; and a good measurement SNR is maintained over all frequencies and sound source directions.

This process opens up the possibility of performing studies that would not have been possible before. By passing through the digital domain, pinna geometry can be altered before measurement in ways not possible with either live subjects or physical castings. For example, we can investigate the effects of individual anthropometric features on the HRTF while holding all other features constant. If human volunteers were relied upon for such a study, the necessary sample size would be infeasibly large in order to compensate for the effects of the many additional parameters that define the shape of the pinna.

Acknowledgments

A preliminary version of this research was presented in a poster session at the 182nd Meeting of the Acoustical Society of America. The authors wish to thank the society members for their encouragement and valuable feedback which helped to refine our initial methodology into that presented here. Hafliði Ásgeirsson, resident technician at the Acoustic and Tactile Engineering laboratory (<https://acute.hi.is>), provided extensive support for the experimental equipment which proved invaluable in the completion of this research. A previous version of this manuscript was reviewed by the publisher's automated English language tool for clarity and the authors have elected to incorporate many of its suggestions.

Authors contributions

The contributions of each author, according to the Contributor Role Taxonomy (CRediT; <https://credit.niso.org/>), are as follows:

- Eric Michael Sumner: Investigation, Methodology, Software, Validation, Visualization, and Writing - Original Draft
- Morris Riedel: Funding Acquisition, Supervision, and Writing - Review & Editing.
- Runar Unnthorsson: Conceptualization, Funding Acquisition, Resources, Supervision, and Writing - Review & Editing.

All authors have read and approved the final version of this manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was performed in the Center of Excellence (CoE) Research on AI-and Simulation-Based Engineering at Exascale (RAISE) receiving funding from EU's Horizon 2020 Research and Innovation Framework Programme [H2020-INFRAEDI-2019-1] under grant agreement [no. 951733]. This work is co-financed by the EUROCC2 project funded by the European High-Performance Computing Joint Undertaking (JU) and EU/EEA states under grant agreement [No 101101903]. Parts of the work have been also supported by the European Digital Innovation Hub (EDIH) of Iceland (EDIH-IS) funded in parts by the Digital Europe Programme. The work described in this paper received financial support from NordForsk's Nordic University Hub 'Nordic Sound and Music Computing Network-NordicSMC' (project no. 86892). This work was supported by the Icelandic Research Fund (IRF), grant [number 2410434-051].

About the authors

Eric Michael Sumner, M. Sc., is a doctoral candidate at the University of Iceland. His research focus is developing hybrid digital/physical data collection processes for training acoustic machine learning models.

Morris Riedel, PhD, is a Professor at the University of Iceland, specialising in parallel and scalable techniques for scientific applications leveraging High-Performance Computing (HPC). His research focuses on parallel Artificial Intelligence (AI) methods enabling scalable machine/deep learning models in application sectors such as remote sensing and healthcare. Still, his research interest also includes the design of HPC algorithms in application domains, such as physical applications, using numerical methods based on known physical laws. He has been actively participating in international projects in HPC/AI for over 20 years and has published extensively in peer-reviewed journals.

Runar Unnthorsson, PhD, is a Professor at the University of Iceland specializing in experimental engineering and product development. His research spans acoustics, non-destructive testing, energy systems, battery technology, and sensory substitution devices, with a focus on sustainable and impactful engineering solutions. He has been actively participating in international and domestic projects for over 25 years and, during this time, has published extensively in peer-reviewed journals.

ORCID

Eric Michael Sumner  <http://orcid.org/0000-0002-6439-9757>

Morris Riedel  <http://orcid.org/0000-0003-1810-9330>

Runar Unnthorsson  <http://orcid.org/0000-0002-1960-0263>

Data availability statement

The data that support the findings of this study are available from the corresponding author, EMS, upon reasonable request.

References

- Ajmera, H., & Gonen, B. (2020). *Virtual reality in health care* [Paper presentation]. 2020 International Conference on Computing, Networking and Communications (ICNC), Big Island, Hawaii (pp. 51–55). <https://doi.org/10.1109/ICNC47757.2020.9049769>
- Algazi, V., Duda, R., Thompson, D., & Avendano, C. (2001). *The CIPIC HRTF database* [Paper presentation]. Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No.01TH8575), New Paltz, NY (pp. 99–102). <https://doi.org/10.1109/ASPAA.2001.969552>
- Andreasen, A., Geronazzo, M., Nilsson, N. C., Zovnercuka, J., Konovalov, K., & Serafin, S. (2019). Auditory feedback for navigation with echoes in virtual environments: Training procedure and orientation strategies. *IEEE Transactions on Visualization and Computer Graphics*, 25(5), 1876–1886. <https://doi.org/10.1109/TVCG.2019.2898787>
- ASTM D2240-00. (2017). *Standard test method for rubber property—durometer hardness*. ASTM International.
- Baumgartner, R., Majdak, P., & Laback, B. August (2014). Modeling sound-source localization in sagittal planes for human listeners. *The Journal of the Acoustical Society of America*, 136(2), 791–802. <https://doi.org/10.1121/1.4887447>
- Bocher, M. (1906). Introduction to the theory of Fourier's series. *The Annals of Mathematics*, 7(3), 81–152. <https://doi.org/10.2307/1967238>
- Bolaños, J. G., & Pulkki, V. (2012). HRIR database with measured actual source direction data. *Audio Engineering Society Convention* 133. <https://www.aes.org/e-lib/browse.cfm?elib=16501>
- Brinkmann, F.. (2019). *Binaural processing for the evaluation of acoustical environments* [PhD thesis]. Technische Universität.
- Bronkhorst, A. (1995). Localization of real and virtual sound sources. *The Journal of the Acoustical Society of America*, 98(5), 2542–2553. <https://doi.org/10.1121/1.413219>
- Burkhard, M. D., & Sachs, R. M. (1975). Anthropometric manikin for acoustic research. *The Journal of the Acoustical Society of America*, 58(1), 214–222. <https://doi.org/10.1121/1.380648>
- Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, 19(2), 272–309. <https://doi.org/10.1080/15213269.2015.1015740>
- Denk, F., Ernst, S. M. A., Ewert, S. D., & Kollmeier, B. (2018). Adapting hearing devices to the individual ear acoustics: Database and target response correction functions for various device styles. *Trends in Hearing*, 22, 2331216518779313. <https://doi.org/10.1177/2331216518779313>
- Ferree, C. E., & Collins, R. (1911). An experimental demonstration of the binaural ratio as a factor in auditory localization. *The American Journal of Psychology*, 22(2), 250–297. <https://doi.org/10.2307/1412921>
- Ghorbal, S., Auclair, T., Soladie, C., & Segquier, R. (2017). *Pinna morphological parameters influencing HRTF sets* [Paper presentation]. Proceedings of the 20th International Conference on Digital Audio Effects (DAFx-17), Edinburgh, UK (pp 1–7).
- Gray, H., & Lewis, W. H. (1918). *Anatomy of the human body* (20th ed.). Lea & Febiger. www.bartleby.com/107/.
- Gray, R., Buzo, A., Gray, A., & Matsuyama, Y. (1980). Distortion measures for speech processing. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28(4), 367–376. <https://doi.org/10.1109/TASSP.1980.1163421>
- Grijalva, F., Martini, L., Florencio, D., & Goldenstein, S. (2016). A manifold learning approach for personalizing HRTFs from anthropometric features. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 24(3), 559–570. <https://doi.org/10.1109/TASLP.2016.2517565>
- Guezenoc, C., & Séguier, R. 06 (2020). A wide dataset of ear shapes and pinna-related transfer functions generated by random ear drawings. *The Journal of the Acoustical Society of America*, 147(6), 4087–4096. <https://doi.org/10.1121/10.0001461>
- Hebrank, J., & Wright, D. (1974). Spectral cues used in the localization of sound sources on the median plane. *The Journal of the Acoustical Society of America*, 56(6), 1829–1834. <https://doi.org/10.1121/1.1903520>
- Humphries, S., Joshi, A., Webb, W. R., & Kanegaonkar, R. Feb (2022). Auricular reconstruction: Where are we now? a critical literature review. *European Archives of Oto-Rhino-Laryngology*, 279(2), 541–556. <https://doi.org/10.1007/s00405-021-06903-5>
- ISO 26101. (2012). *Acoustics. Test methods for the qualification of free-field environments*. International Organization for Standardization.
- Jenny, C., & Reuter, C. (2020). Usability of individualized head-related transfer functions in virtual reality: Empirical study with perceptual attributes in sagittal plane sound localization. *JMIR Serious Games*, 8(3), e17576. <https://doi.org/10.2196/17576>
- Katz, B. F. G. 11 (2001). Boundary element method calculation of individual head-related transfer function. I. Rigid model calculation. *The Journal of the Acoustical Society of America*, 110(5 Pt 1), 2440–2448. <https://doi.org/10.1121/1.1412440>

- Liu, P. R., & Meng, M. Q. H. (2007). *Sensory substitution for internet-based e-healthcare robotic system* [Paper presentation]. 2007 IEEE International Conference on Robotics and Biomimetics (ROBIO), Sanya, China (p. 31–36). <https://doi.org/10.1109/ROBIO.2007.4522130>
- Morimoto, M., & Ando, Y. (1980). On the simulation of sound localization. *Journal of the Acoustical Society of Japan (E)*, 1(3), 167–174. <https://doi.org/10.1250/ast.1.167>
- Pelzer, R., Dinakaran, M., Brinkmann, F., Lepa, S., Grosche, P., & Weinzierl, S. (2020). Head-related transfer function recommendation based on perceptual similarities and anthropometric features. *The Journal of the Acoustical Society of America*, 148(6), 3809–3817. <https://doi.org/10.1121/10.0002884>
- Rayleigh, L. (1909). On the perception of the direction of sound. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 83(559), 61–64. <https://doi.org/10.2307/93054>
- Rosati, G., Oscari, F., Reinkensmeyer, D. J., Secoli, R., Avanzini, F., Spagnol, S., & Masiero, S. (2011). *Improving robotics for neurorehabilitation: Enhancing engagement, performance, and learning with auditory feedback* [Paper presentation]. 2011 IEEE International Conference on Rehabilitation Robotics, Zurich (pp. 1–6). <https://doi.org/10.1109/ICORR.2011.5975373>
- Spagnol, S., Geronazzo, M., & Avanzini, F. (2013). On the relation between pinna reflection patterns and head-related transfer function features. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(3), 508–519. <https://doi.org/10.1109/TASL.2012.2227730>
- Spagnol, S., Hoffmann, R., Herrera Martínez, M., & Unnthorsson, R. (2018). Blind wayfinding with physically-based liquid sounds. *International Journal of Human-Computer Studies*, 115, 9–19. <https://doi.org/10.1016/j.ijhcs.2018.02.002>
- Spagnol, S., Miccini, R., & Unnthorsson, R. (2020). The Viking HRTF. dataset v2. <https://doi.org/10.5281/zenodo.4160401>
- Spagnol, S., Miccini, R., Onofrei, M. G., Unnthorsson, R., & Serafin, S. (2021). Estimation of spectral notches from pinna meshes: Insights from a simple computational model. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 29, 2683–2695. <https://doi.org/10.1109/TASLP.2021.3101928>
- Spagnol, S., Wersényi, G., Bujacz, M., Bălan, O., Herrera Martínez, M., Moldoveanu, A., & Unnthorsson, R. (2018). Current use and future perspectives of spatial audio technologies in electronic travel aids. *Wireless Communications and Mobile Computing*, 2018(1), 3918284. <https://doi.org/10.1155/2018/3918284>
- Sumner, E. M., Unnthorsson, R., & Riedel, M. (2022). *Replicating human sound localization with a multi-layer perceptron* [Paper presentation]. Proceedings of the 19th Sound and Music Computing Conference, Saint-Étienne. <https://doi.org/10.5281/zenodo.6797854>
- Wenninger, M. J. (1979). Introduction to geodesic symbolism and classification. In *Spherical models* (pp. 98–100). Cambridge University Press.